

Brownian Dynamics Studies on DNA Gel Electrophoresis. I.

Numerical Method and Quasi-Periodic Behavior of Elongation-Contraction Motions

Ryuzo Azuma* and Hajime Takayama

Institute for Solid State Physics, The university of Tokyo

5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8581, Japan

(Dated: February 1, 2008)

Abstract

Dynamics of individual DNA undergoing constant field gel electrophoresis (CFGE) is studied by a Brownian dynamics (BD) simulation method which we have developed. The method simulates electrophoresis of DNA in a 3 dimensional (3D) space by a chain of electrolyte beads of hard spheres. Under the constraint that the separation of each pair of bonded beads is restricted to be less than a certain fixed value, as well as with the excluded volume effect, the Langevin equation of motion for the beads is solved by means of the Lagrangian multiplier method. The resultant mobilities, μ , as a function of the electric field coincide satisfactorily with the corresponding experimental results, once the time, the length and the field of the simulation are properly scaled. In relatively strong fields quasi-periodic behavior is found in the chain dynamics, and is examined through the time evolution of the radius of the longer principal axis, $R_l(t)$. It is found that the mean width of a peak in $R_l(t)$, or a period of one elongation-contraction process of the chain, is proportional to the number of beads in the chain, M , while the mean period between two such adjacent peaks is proportional to M^0 for large M . These results, combined with the observation that the chain moves to the field direction by the distance proportional to M in each elongation-contraction motion, yield $\mu \propto M^0$. This explains why CFGE cannot separate DNA according to their size L ($\propto M$) for large L .

I. INTRODUCTION

Gel electrophoresis is a most widely used technique to separate DNA according to their length L under an electric field.¹ The constant field method is quite efficient for DNA of small sizes. It is well known, however, that the separation does not work well for DNA of sizes longer than a certain value which depends on the field and the “pore” diameter of the gel.² The non-constant field methods such as the pulsed-field gel electrophoresis³ have been empirically invented and they have remarkably improved the shortcoming. Nevertheless, the understandings on these gel electrophoresis methods, even on the constant field method, still remain unsatisfactory from the view point of statistical physics of polymers.

By constant field gel electrophoresis (CFGE), DNA molecules with the radius of inertia, R_I , smaller than the pore diameter of gel, a_{gel} , are *sieved* by gel.^{4,5} In this sieving regime DNA with the smaller R_I migrate the faster through the gel. In the opposite case with $R_I \gtrsim a_{\text{gel}}$, DNA is considered to move through a “primitive chain”⁶ which is determined by stochastic movement of its front end in pores of the gel under the field.^{7,8,9} This picture, denoted as the biased reptation model, successfully describes various features of CFGE, such as the L -dependence of the migration mobility, μ , observed experimentally¹⁰ in sufficiently weak fields. More recently, the model of the biased reptation with fluctuations has been proposed¹¹ which well describes the observed field dependence of μ in moderate fields.¹ When the field is further increased, the field-dependence of μ begins to be saturated. In this regime, interesting behavior of an individual DNA has been observed: “hernias”-like configurations by the computer simulations^{12,13} and “quasi-periodic” behavior, i.e., “elongation-contraction” motion in a cyclic way by the real-time fluorescent microscopy experiments.^{14,15} Indeed, the latter experimental method combined with optical or magnetic tweezers for a molecule reveals that the elasticity of DNA is originated from the entropy associated with its static conformations.^{16,17,18,19} However, dynamic properties of DNA under CFGE are not described in a unified way by this picture alone.

In order to get further insights into the problem mentioned above, we have developed a Brownian dynamics (BD) method which simulates electrophoresis of DNA in a 3 dimensional (3D) space. The method has been extensively applied to the CFGE, in particular, the elongation-contraction motion of the chain under CFGE, the results of which we will discuss in the present series of papers. In the present paper I our numerical method is explained and

global aspects, such as the mean mobility μ and time evolutions of the velocity of the center of mass, $v_G(t)$, and of the radius of the longer principal axis, $R_l(t)$, are investigated. In the accompanying paper II²⁰ we analyze dynamics of ‘defects’ introduced by de Gennes²¹ in a chain under CFGE. Finally in paper III²² we will examine conformational changes of the coarse-grained chain under CFGE in details, and will argue that the field-parallel component of its elongation-contraction motion is essentially understood as a deterministic dynamics of an elastic string in a 1D space with an obstacle, thereby an origin of the elasticity will be shown to be the conformational entropy of the chain.

In the Brownian dynamics (BD) simulations extensively performed in the present series of studies, DNA is modeled as a chain of spherical electrolyte beads with a constant radius $1/2$ in a 3D space filled with a solvent. Each bead interacts with other beads of the chain and spatially-fixed beads constituting gel by hard-core repulsive interactions. It also interacts with its two bonded beads by a non-linear elastic force which keeps the distance between the two to be less than a threshold value which we put $\sqrt{2}$. To solve the Langevin equation of motion under these constraints as well as under the field and random forces, we employ a BD algorithm with the Lagrangian multiplier method. It is similar to the one by Deutsch and Madden,^{12,13} but is essentially different from their method in the following respects. In their method, the time pitch to solve the equation is adjusted so as to keep a fixed distance between bonded beads under the tensile forces which are evaluated by the Lagrangian multiplier method. In our method, on the other hand, the time pitch is fixed at a certain value, which we regard as the mean period of random forces. Within each time step we evaluate paths of all beads affected by the constraints, i.e., by hard-core interactions and extremely nonlinear elastic interactions between a pair of beads. These effects are reread as the supposed constraining forces with the appropriate Lagrangian multipliers in an iterative way. The resultant constraining forces thus specified yield a new set of bead positions without any violation to the constraints. To our knowledge the present BD analysis is the first extensive simulation of DNA gel electrophoresis in a 3D space based on the Langevin equation involving purely mechanical, microscopic forces which are supposed to act on monomers of DNA.

The purposes of the present paper I are to explain our BD method and to demonstrate its results on some global properties of DNA under CFGE. The resultant mobility μ as a function of the field turns out to agree well with those observed experimentally once a set of

three parameters, l^* , τ^* and E^* , are properly specified. They relate respectively the scales of length, time and field strength of the simulation and the experiments.

In the BD simulation under relatively large fields quasi-periodic time evolution is observed in $R_l(t)$ (see Fig. 5 below), as well as in $v_G(t)$, as has been observed experimentally.^{14,15} Each local maximum of $R_l(t)$ at t_{\max}^n with $n = 1, 2, \dots$ corresponds to a conformation that a chain just gets rid of trapping due to gel. The local minimum at $t_{\min:l}^n$ preceding this maximum is then considered to be the instance that, being trapped by gel, the chain starts to elongate from a certain coiled conformation, while at time $t_{\min:r}^n$ of the local minimum subsequent to t_{\max}^n the chain retrieves another coiled state. We establish a method to properly classify a whole time sequence of $R_l(t)$ into peak parts, from $t_{\min:l}^n$ to $t_{\min:r}^n$ and the rest. We call the former parts *deterministic* ranges and the latter *stochastic* ones. We then investigate statistics of periods of the deterministic ranges, $D_n \equiv t_{\min:r}^n - t_{\min:l}^n$, and those of the stochastic ranges, $(1/\lambda)_n \equiv t_{\min:l}^{n+1} - t_{\min:r}^n$. It is found that $\{D_n\}$ obey a Gaussian distribution with the mean $D \equiv \overline{D_n} \propto M$ and the width $\Delta \propto D$, while $\{(1/\lambda)_n\}$ do a Poisson distribution with the mean $1/\lambda \equiv \overline{(1/\lambda)_n} \propto M^0$ for large M , where M is the number of beads of the chain and the overline stands for the averages over n , i.e., peak structures. These results, combined with the observation that in each elongation-contraction motion the chain moves to the field direction by the distance proportional to M , yield $\mu \propto M^0$ and explain why CFGE cannot separate DNA according to their sizes L ($\propto M$) for large L .

The organization of the present paper is as follows. In the next section we explain our BD method. The mobility μ of the chain obtained by the BD method is presented and compared with the experimental results in Sec. III. The method of analysis of peak structures in $R_l(t)$ is explained, and statistical nature of the *deterministic* and *stochastic* ranges is discussed in Sec. IV. The last section is devoted to further discussions of our results.

II. THE BROWNIAN DYNAMICS METHOD

A chain of spherical beads in a continuous 3D space, shown schematically in Fig. 1, is investigated by a new type of Brownian dynamics (BD) calculation. The hard core interaction is assumed in order for centers of any pair of beads not to come closer than 1 (the excluded volume effect), where the radius of beads is set $1/2$. Furthermore distances between neighboring beads, l , are restricted to $l < \sqrt{2}$. We suppose that this is realized by

the extremely nonlinear elastic interaction between the neighboring beads: it is infinite for $l > \sqrt{2}$ and zero otherwise. These interactions assure the self-avoiding walk of the chain.

The BD method we adopt in the present work is to solve the following Langevin equation of motion for the chain,

$$\zeta \dot{\mathbf{x}}_i = q \mathbf{E}_b + \sum_{\alpha} \sum_j \mathbf{S}_{ij}^{\alpha} + \sum_{\alpha} [\mathbf{F}_i^{\alpha} - \mathbf{F}_{i-1}^{\alpha}] + \mathbf{f}_i, \quad (1)$$

where \mathbf{x}_i is the position of the center of the i -th bead, and ζ , q , and \mathbf{E}_b are the viscosity, the charge of a bead and the (bare) external field, respectively. The second and the third terms in r.h.s. are the constraining forces required for the beads to satisfy the conditions,

$$\|\mathbf{x}_{ij}\| = \|\mathbf{x}_j - \mathbf{x}_i\| > 1 \quad (2)$$

$$\|\mathbf{l}_i\| = \|\mathbf{x}_{i+1} - \mathbf{x}_i\| < \sqrt{2}. \quad (3)$$

The vectors \mathbf{S}_{ij}^{α} and \mathbf{F}_i^{α} are respectively written as $\theta_{ij}^{\alpha}[\mathbf{x}_i - \mathbf{x}_j]$ and $\phi_i^{\alpha}[\mathbf{x}_i - \mathbf{x}_{i+1}]$ by introducing the Lagrangian multipliers θ_{ij}^{α} and ϕ_i^{α} . The last term in eqn. 1 stands for the random force from a solvent. Its distribution is characterized by the following averages

$$\langle \mathbf{f}_i \rangle = \mathbf{0} \quad (4)$$

$$\langle \mathbf{f}_i(t) \mathbf{f}_j(t') \rangle = 2\zeta k_B T \Delta t \delta_{ij} \delta(t - t'). \quad (5)$$

Here, k_B , T , and Δt are the Boltzmann constant, the absolute temperature, and the average period of the random force, respectively.

To solve the one step evolution of eqn. 1, thereby the terms $\{\mathbf{S}_{ij}^{\alpha}, \mathbf{F}_i^{\alpha}\}$ with the Lagrangian multipliers are determined, we examine the full Langevin equation with the inertia term and the forces arising from the hard core interaction and the nonlinear elastic interaction mentioned above. We solve the equation in a time interval of $\Delta t > t > 0$, and take the overdamped limit, $m/\zeta \rightarrow 0$, where m is the supposed mass of the bead. The constraints of eqns. 2 and 3 are supposed to be satisfied by positions of the beads at $t = 0$, denoted by $\{\mathbf{x}_i^0\}$. At this instance \mathbf{f}_i and $q\mathbf{E}_b\Delta t$ are added as impulsive forces with the integrated strength $\tilde{\mathbf{f}}_i$ where $\tilde{\mathbf{f}}_i = \mathbf{f}_i + q\mathbf{E}_b\Delta t$. When the constraints are discarded, the solution in the overdamped limit is described by the relation between the displacement $\Delta\mathbf{x}_i$ in this time interval and the force as

$$\Delta\mathbf{x}_i = \frac{1}{\zeta} \tilde{\mathbf{f}}_i, \quad (6)$$

which is just what is obtained by simply integrating eqn. 1 without the second and the third terms.

In the case that positions of a pair of beads $\mathbf{x}_i^{01} \equiv \mathbf{x}_i^0 + \Delta\mathbf{x}_i$ and $\mathbf{x}_j^{01} \equiv \mathbf{x}_j^0 + \Delta\mathbf{x}_j$ violate the constraint of eqn. 2, we solve the full Langevin equation for this pair with the hard core scattering, which can be either elastic or inelastic. We then obtain, again in the overdamped limit, positions \mathbf{x}_i^1 and \mathbf{x}_j^1 at $t = \Delta t$ as shown in Fig. 2a in case of the inelastic scattering. The result is reread as the consequence of the constraining force, $\mathbf{S}_{ij}^1 \equiv \theta_{ij}^1[\mathbf{x}_i^1 - \mathbf{x}_j^1]$, which is determined by the conditions $\mathbf{x}_i^1 - \mathbf{x}_i^0 = (\tilde{\mathbf{f}}_i + \mathbf{S}_{ij}^1)/\zeta$ and $\mathbf{x}_j^1 - \mathbf{x}_j^0 = (\tilde{\mathbf{f}}_j - \mathbf{S}_{ij}^1)/\zeta$ (note that $\mathbf{S}_{ji}^1 = -\mathbf{S}_{ij}^1$). This constraining force \mathbf{S}_{ij}^1 just corresponds to the one in the second term of eqn. 1. In the case \mathbf{x}_i^{01} and \mathbf{x}_{i+1}^{01} violate eqn. 3, the nonlinear elastic force is supposed to reduce the distance $\|\mathbf{l}_i\|$ to $\sqrt{2}$. This effect is represented by the constraining force \mathbf{F}_i^1 ($= -\mathbf{F}_{i+1}^1$), which is added to the third term of eqn. 1. Its Lagrangian multiplier ϕ_i^1 is determined in such a way that $\|\mathbf{x}_i^1 - \mathbf{x}_{i+1}^1\| = \sqrt{2}$ with $\mathbf{x}_i^1 = \mathbf{x}_i^0 + \mathbf{F}_i^1\Delta t$ and $\mathbf{x}_{i+1}^1 = \mathbf{x}_{i+1}^0 + \mathbf{F}_{i+1}^1\Delta t$ (see Fig. 2b). If \mathbf{x}_i^{01} involves violation to more than one constraints of eqns. 2 and/or 3, the above-mentioned procedures are applied to evaluate the constraining forces for each pair of beads with the common \mathbf{x}_i^{01} . If, on the other hand, \mathbf{x}_i^{01} does not involve any violation to eqns. 2 and 3, we put $\mathbf{x}_i^1 = \mathbf{x}_i^{01}$.

In the configuration $\{\mathbf{x}_i^1\}$, it is generally expected that new pairs of beads may violate the constraints since we have tried to remedy the violations in $\{\mathbf{x}_i^{01}\}$ in a pairwise way. Then we repeat the above-mentioned procedure, in which $\tilde{\mathbf{f}}_i$ is replaced by $\tilde{\mathbf{f}}_i + \mathbf{C}_i^1$, where $\mathbf{C}_i^1 = \sum_j \mathbf{S}_{ij}^1 + \mathbf{F}_i^1 - \mathbf{F}_{i-1}^1$ is the constraining force determined by the first procedure. If a configuration without any violation to the constraints is obtained by the α_{con} times repetition of this procedure, we regard $\{\mathbf{x}_i^{\alpha_{\text{con}}}\}$ as the solution of eqn. 1 evolved from $\{\mathbf{x}_i^0\}$ by one unit of time Δt . In this manner the multi-scattering processes in the interval of Δt , including those due to the nonlinear elastic interaction, can be taken into account, and the corresponding forces $\{\mathbf{S}_{ij}^\alpha, \mathbf{F}_i^\alpha\}$ in eqn. 1 are determined.

Although our BD method is based on the overdamped Langevin equation with the Lagrangian multipliers, it distinctly differs from the one due to Deutsch and Madden.^{12,13} In the latter the equidistant condition on all the pairs of beads is imposed, and the constraining tensile forces associated with it are evaluated at the instance when random forces are applied. Then the time interval Δt is adjusted so that the configuration at $t + \Delta t$ in fact obeys the equidistant condition within accuracy of 0.1%. In contrast, our BD method takes

into account a similar, but looser, constraint on the bond length as well as the excluded volume effect of the beads by explicitly examining the supposed multiple scattering within the interval Δt which is fixed. Therefore there is no finer time unit other than Δt in solutions of our BD method. It is the mean period of random forces $\{\mathbf{f}_i\}$.

In the present work we simulate dynamics of a chain in a continuous 3D space confined into a cubic box with a volume 360^3 . The periodic boundary condition is imposed to each direction of the box. The gel is represented by a network of immobile bars arranged so that they form a simple cubic lattice. Each bar consists of tightly connected beads with a radius $1/2$, and its interaction with a bead of the mobile chain is computed according to the constraint of eqn. 2. We employed three lattices of gel with different lattice distances $a_{\text{gel}} = 10, 17$, and 20 , which are considered to correspond to the 'pore' diameter of agarose gel.² The lengths of chains, M , used in the BD calculation are $M = 30, 40, 60, 80, 160, 240$, and 320 . We have performed typically 64 runs whose initial configurations are given independently. The direction of the field is fixed at $(1, 1, 1)$. The time is measured in unit of Δt , i.e., 1 BD step.

Here we want to emphasize that the present BD method correctly reproduces the expected static and dynamic properties of a real (not phantom) polymer in a 3D space under the vanishing field: $R_I \propto M^{\nu_F}$ with $\nu_F \simeq 0.59 \pm 0.01$ and the Einstein relation $D_G \propto \mu/M$ are confirmed (see Fig. 3 in Sec. III). Here, ν_F , R_I , and D_G are the Flory exponent, the radius of inertia, and the diffusion constant of the center of mass, respectively.

The number of the multiple scattering, α_{con} above mentioned, is expected to be at most of the order of M^2 . It has turned out that α_{con} becomes relatively larger when the chain is in an extended configuration under gel electrophoresis. We quote here typical figures for a chain with $M = 80$: $\alpha_{\text{con}} \simeq 4 \times 10^1, 5 \times 10^2$ for $E = 0, 0.032$, respectively.

III. MOBILITY

The most basic quantity that has been commonly measured in CFGE of DNA is the long time average of velocity divided by the electric field, or the mobility μ . In the present work, we have implemented long time BD simulations with a highly efficient algorithm using a massive parallel computer and have eventually obtained μ , hereafter denoted by μ_{BD} , with enough statistical precision for all the set of parameters investigated. Here we only mention

typical figures of computation to get μ_{BD} of a $M = 240$ chain for each run: more than 150 hours are consumed on a single processor of SGI2800 system to pass 10^6 BD steps. In the present work the following values of the parameters are chosen: $\Delta t = 1$, $\zeta = 10$ and $k_{\text{B}}T = 2$. The results are shown in Fig. 3 with the experimental data due to Heller *et al.*²³ Here and hereafter the value of field E stands for qE_{b} of eqn. 1.

The simulated μ_{BD} at $E = 0$ in Fig. 3 is evaluated through the Einstein relation

$$\frac{\mu_{\text{BD}}^0}{M} = D_{\text{G}} \equiv \lim_{t \rightarrow \infty} \frac{1}{6t} \langle (\mathbf{R}_{\text{G}}(t) - \mathbf{R}_{\text{G}}(0))^2 \rangle|_{E=0}, \quad (7)$$

where $\mathbf{R}_{\text{G}}(t)$ is the position of the center of mass of the chain. The diffusion constant D_{G} in the r.h.s. is evaluated by the BD simulation in a vanishing field in the time range $t = (5 \sim 10) \times 10^6$. The resultant μ_{BD}^0 turns out to be compatible with the value of $\lim_{E \rightarrow 0} \mu_{\text{BD}}$ as seen in the figure. This confirms that the Einstein relation in fact holds in the present BD simulation for the chain.

In order to compare the BD results with the experimentally observed data, we have to specify proper conversion of values of the parameters in the simulation and the experiment. Let us first consider the length scale of a chain. In the accompanying paper II we examine in detail the embodiment of ‘defects’ due to de Gennes²¹ in our BD chain and obtain the mean distance $\langle n \rangle$ of adjacent ‘defects’ along the chain under $E = 0$ as $\langle n \rangle \simeq 5.7$. This value, in unit of the mean distance between neighboring beads, $\langle l \rangle$, is almost independent of the set of parameters a_{gel} and M we have examined. Regarding $\langle n \rangle$ as an estimate of the ‘persistent length’ in the BD simulation, \mathcal{P}_{BD} , we put $\mathcal{P}_{\text{BD}} = c\langle n \rangle$, where c is an adjustable parameter whose value is nearly equal to unity (note that $\sqrt{2} \geq \langle l \rangle \geq 1$). The conversion factor of the length-scale l^* is then determined by

$$\mathcal{P}_{\text{exp}} = l^* \mathcal{P}_{\text{BD}} = cl^* \langle n \rangle, \quad (8)$$

where \mathcal{P}_{exp} is the persistence length experimentally measured. With $\mathcal{P}_{\text{exp}} \sim 60\text{nm}$ at $T_{\text{exp}} \sim 300\text{K}$ taken from Ref. 24, we obtain $cl^* \sim 10\text{nm}$.

Another length scale of importance is the mean pore size of gel, which is rather difficult to be accurately specified even in the experiments.¹ Here we simply suppose that DNA in 1% agarose corresponds to a BD chain in the gel with $a_{\text{gel}} = 17$. With \mathcal{P}_{BD} above fixed, this value of a_{gel} is about three times larger than that of \mathcal{P}_{BD} .

Next let us relate E in the BD simulation to the experimental field, E_{exp} , in V/cm. For this purpose we impose that the ratios $\Delta\varepsilon/k_{\text{B}}T$ in the BD calculation and the experiment

be identical, where $\Delta\varepsilon$ is the energy that a part of chain of a length equal to the persistence length \mathcal{P} gains when it moves a distance of \mathcal{P} under the electric field:

$$\frac{\sigma E_{\text{exp}} \mathcal{P}_{\text{exp}}^2}{k_{\text{B}} T_{\text{exp}}} = \frac{E \mathcal{P}_{\text{BD}}^2}{k_{\text{B}} T_{\text{BD}}}, \quad (9)$$

or

$$E_{\text{exp}} = E^* E; \quad E^* = \frac{k_{\text{B}} T_{\text{exp}} \mathcal{P}_{\text{BD}}^2}{2\sigma \mathcal{P}_{\text{exp}}^2} \simeq 5 \times 10^2 c^{-2} \text{ V/cm}. \quad (10)$$

where σ is the density of charge of DNA, and $k_{\text{B}} T_{\text{BD}} = 2$ has been used. Quoting further the experimental value of $\sigma \sim 0.6e^-/\text{\AA}$,²⁴ as well as $cl^* \sim 10\text{nm}$ at $T_{\text{exp}} \sim 300\text{K}$ in eqn. 8, we obtain the last expression for E^* in the above equation. As for the ordinates of Fig. 3 we may relate μ_{BD} to $\mu_{\text{exp}}[\text{cm}^2/\text{V} \cdot \text{s}]$ by

$$\mu_{\text{exp}} = \mu^* \mu_{\text{BD}}; \quad \mu^* = \frac{l^*}{\tau^* E^*} \simeq 2 \times 10^{-9} \frac{c}{\tau^*} [\text{cm}^2/\text{V} \cdot \text{s}], \quad (11)$$

where $\tau^*[\text{s}]$ is the real time which corresponds to one discretized time step Δt in the BD calculation with $\zeta = 10$ (or $\Delta t/\zeta = 0.1$).

The two sets of data of μ in Fig. 3 are drawn as follows. The data μ_{exp} are plotted directly using the units indicated on the right ordinate and the upper abscissa, whereas the units for μ_{BD} are indicated on the left ordinate and the lower abscissa with $c^2 = 1.60$ and 1.26 respectively for $M = 240$ and 320 . The scales of the two ordinates are related by means of $\tau^* = 1.4 \times 10^{-6}$. We see from the figure that the field dependence of μ_{BD} , particularly that of $M = 240$, well coincide with that of μ_{exp} of DNA with the corresponding length we have specified, i.e., 4.3 and 6.5 kbp DNA for $M = 240$ and 320 , respectively. This is quite satisfactory if we take into account the fact that only three adjustable parameters are used to draw the two sets of data in Fig. 3. We therefore consider that the present BD model catches up even quantitatively essences of DNA dynamics under CFGE.

When we compare μ_{BD} with μ_{exp} further in detail, however, there are some unsatisfactory features in the BD results. As compared with μ_{exp} of 6.5 kbp DNA, an E -independent branch of μ_{BD} has not been ascertained in the weak field limit for the $M = 320$ chain. One possible reason for this may be the insufficient time of simulation (up to $t = 10^6$ in finite fields) for this length of chain in weak fields. Thus, within the limited CPU time, it is a rather hard task for the present BD simulation to judge the field dependence of μ_{BD} , proportional to E or E^2 , in weak fields.^{25,26,27} In the strong field regime, on the other hand, μ_{BD} tend to

saturate, while, μ_{exp} are still significantly increasing with E . These different tendencies are seen more clearly if we compare μ_{exp} measured in 10V/cm with the corresponding μ_{BD} (not seen). The reason of this discrepancy is not clear at the moment.

We have also examined μ_{BD} of the chains in gels with $a_{\text{gel}} = 20$ and 10. If only the volume ratio of gel is considered, they correspond roughly to 0.75% and 3% agaroses, respectively. Qualitatively, μ_{BD} becomes larger and its E -dependence becomes weaker for the larger a_{gel} as expected, but these tendencies are quantitatively weaker than those observed experimentally. The origins of the discrepancy may be attributed to our model for gel, i.e., a perfectly rigid, regular jungle gym.

IV. ELONGATION-CONTRACTION MOTION

A. Motion of an individual chain

In fields larger than a certain value depending on M and a_{gel} , chains in the present 3D BD simulation exhibit quasi-periodic behavior, i.e., they exhibit elongated and contracted shapes alternatively, as observed in the experiments^{14,15,28} and in the previous 2D BD models.^{12,29,30} It is here emphasized that such quasi-periodic behavior has been observed within the field range where we have discussed the E -dependence of μ in Fig. 3. We show a typical conformational change of a chain observed in one MD run of CFGE with the parameters $M = 240$, $E = 0.032$, and $a_{\text{gel}} = 17$ in Fig. 4. Within a time window of the figure the elongation-contraction motion occurs three times at around $t = 2, 6$ and 11×10^5 . At ranges between them centered around $t = 4$ and 8×10^5 a chain is in a rather compact form. In this range it can happen, though not so frequently, that the front and rear ends exchange their roles as seen around $t = 9 \times 10^5$ in the figure. It is also pointed out here that the front end of a chain moves in the field direction with almost a constant rate.

B. Quasi-periodic behavior of $R_l(t)$ and $v_G(t)$

In Fig. 5 we show the time evolution of the radius of longer principal axis, $R_l(t)$, and the velocity of the center of mass, $v_G(t)$, in the MD run whose chain conformational change is shown in Fig. 4. The fluctuation in $R_l(t)$ is notable, showing successive Λ -shaped peaks.

We also note that just before the maxima of $R_l(t)$, minima are observed in $v_G(t)$. These features are qualitatively in good agreement with those observed in the experiment.¹⁵

The time evolution of $R_l(t)$ and $v_G(t)$ in Fig. 5 looks quasi-periodic. Actually it was found by the experiment that the autocorrelation function of $R_l(t)$, $C_{RR}(t)$, exhibits a damped oscillation.¹⁵ This is also the case for our present BD data. In Fig. 6 $C_{RR}(t)$ evaluated for various M are presented. For $M \geq 80$, $C_{RR}(t)$ exhibit the first undershoot below the line of $C_{RR}(t) = 0$, and even the first overshoot is clearly seen for $M \geq 160$. From these results we may conclude that, in chains with $M \geq 160$ under $E = 0.032$, elongation-contraction motions occur quasi-periodically. In the inset of Fig. 6 we show nearly linear dependence of the period of oscillation $\tau = 4t_0$ on M , where t_0 is defined as the time of the first intercept of $C_{RR}(t)$ with the abscissa.¹⁵

C. Analyses on peak structures of $R_l(t)$

Oana *et al.*¹⁵ argued that the steady-state time evolution of DNA under CFGE is classified into two types of behavior. One is the elongation-contraction motion which looks apparently deterministic. We call the time branches in which such a motion is undergoing the *deterministic* ones (ranges centered at around $t = 2, 6$ and 11×10^5 in Fig. 4). The other is what is observed between two deterministic branches. We call them *stochastic* branches. Oana *et al.* approximately described the quasi-periodic behavior of $R_l(t)$ in terms of a simplified function: peaks in $R_l(t)$ are approximated by Λ -shape branches whose widths are assumed to obey a Gaussian distribution, and others by constants whose duration times are assumed to obey a Poisson distribution. They calculated $C_{RR}(t)$ by means of this model function for various cases and remarked that when λD becomes larger than unity a damped oscillation in $C_{RR}(t)$ becomes prominent. Here $1/\lambda$ and D stand for the mean period of the stochastic branches and the mean width of the deterministic ones, respectively.

Following the classification due to Oana *et al.*, we analyze our BD results by making use of an algorithm which appropriately picks up peaks of $R_l(t)$. For this purpose we first pick up time sets $\{t_{\min:l}^i, t_{\max}^i, t_{\min:r}^i\}$, i.e., the times of preceding local minimum of the i -th maximum, the i -th maximum itself, and the subsequent minimum. Then we select out peaks whose height relative to their adjacent minima is larger than a certain threshold value. The latter is chosen in such a way that the sum of periods of peak branches selected surmounts

more than 50% of the total duration of the observation. This procedure divides a whole time sequence of $R_i(t)$ into the two types of branches whose numbers are equal to each other. It picks up ensemble of various peaks but not a limited numbers of peaks of special shapes. The number of peaks thus selected is more than a hundred for each set of parameters.

D. Statistics of deterministic and stochastic branches

For the set of times $\{t_{\min:l}^n, t_{\max}^n, t_{\min:r}^n\}$ of peaks picked up by the procedure described above we examine the distribution of peak widths (or periods of deterministic branch) $D_n \equiv t_{\min:r}^n - t_{\min:l}^n$ and that of duration times between neighboring peaks (periods of stochastic branch) $1/\lambda_n \equiv t_{\min:r}^n - t_{\min:l}^{n-1}$. The results for $M = 240, E = 0.032$ and $a_{\text{gel}} = 17$ are shown in Fig. 7. They numerically confirm the basic assumptions of Oana *et al.*, i.e., $\{D_n\}$ nearly obey a Gaussian distribution, while $\{1/\lambda_n\}$ do a Poisson distribution. The mean and the variance of the former are given by $D = 3.3 \times 10^5$ and $\Delta \simeq 0.2D$, respectively, while the mean width of the latter is given by $\lambda D \simeq 1.7$. The sum $D + 1/\lambda$ should be equal to $\tau = 4t_0$ evaluated from $C_{\text{RR}}(t)$ in Sec. IV B, which is verified within our numerical accuracy. The ratio $\lambda D \simeq 1.6$ fulfills the condition for a damped oscillation to be observed in $C_{\text{RR}}(t)$.

The ratio λD is found to decrease below unity for M which is smaller than a certain value. In Fig. 8, D and $1/\lambda$ are plotted against M for $E = 0.032$ and $a_{\text{gel}} = 17$. The results strongly suggest that $D \propto M^1$ and $1/\lambda \propto M^0$ for large M . In Fig. 9, the M -dependence of ratios λD and Δ/D are presented. As expected from the results shown in Fig. 8, λD is an approximately linearly increasing function for large M . On the other hand, Δ/D is almost constant (~ 0.3), indicating that the Gaussian distribution of $\{D_n\}$ is scaled by the mean D alone. It should be noted that, for $M \gtrsim 100$ where $\lambda D \gtrsim 1$ holds, $C_{\text{RR}}(t)$ clearly exhibits a damped oscillation as has been observed in Fig. 6. Also the result $4t_0 \propto M$ for $M \gtrsim 100$ shown in the inset of Fig. 6 is in accordance with the fact that λD is larger than unity for such M . For sufficiently large M the deterministic branches dominate a whole sequence of the time evolution, and so $4t_0 \simeq D \propto M$ holds. These results consistently indicate that there exists a crossover between the regimes with and without elongation-contraction motions at around $M \cong 100$ in CFGE with the parameters studied, i.e., $E = 0.032$ and $a_{\text{gel}} = 17$.

We have also examined the field dependence of D and $1/\lambda$ for a fixed $M (= 240)$, the results of which are shown in Fig. 10. In strong fields $E \gtrsim 0.06$, DE becomes to be saturated,

while E/λ decreases with increasing E . As a consequence, $D\lambda$ is an increasing function of E , indicating that the deterministic branches, or in other words, the elongation-contraction motions survive and even dominate a whole evolution of the chain in sufficiently strong fields.

Lastly, in Fig. 11, the averaged time evolution of individual peaks of $R_l(t)$ are plotted for various M and E . The abscissa and ordinate are normalized by the period $\tau \equiv 4t_0$ and the averaged peak height, respectively. It is seen that with these normalizations, the data with different M and E almost lie on a universal curve during the period from -0.6τ to 0.4τ . The results obtained above such as $D \propto M, E^{-1}$, $\Delta \propto D(\propto M)$ can be derived from the scaling behavior of $R_l(t)$ in Fig. 11. Such scaling behavior of $R_l(t)$ including the fact that the normalized peak exhibits an anti-symmetric shape is just what was observed experimentally.¹⁵ It was mentioned in Ref. 15 that the ratio of slope of the Λ -shape immediately after the maximum to that before it is approximately constant (~ 3) for all conditions they employed. The corresponding ratio of the BD result is ~ 2 . The origin of this quantitative difference is not clear at the moment.

V. DISCUSSIONS

We have performed extensive simulations on the DNA constant-field gel electrophoresis (CFGE) by means of a new Brownian dynamics (BD) method which we have recently developed. Our BD method is based on a *generalized* bead-spring model for a polymer in a 3D space with rigid, regularly arrayed obstacles. By the word ‘*generalized*’ we mean that the ‘spring force’ is not linear but extremely nonlinear: it is infinite when the distance l between two beads connected by the ‘spring’ is larger than $\sqrt{2}$, while it is zero when $l < \sqrt{2}$. Also in the model we strictly take into account of the excluded volume effects between any pair of beads whose diameter is put unity. These forces are supposed to be of a microscopic origin, i.e., molecular forces acting on monomers of a polymer. It should be emphasized that there involves no characteristic scale in these interactions except for the maximum distance between the connected beads ($\sqrt{2}$) relative to the bead diameter (unity).

In a vanishing field the viscosity ζ (or $\Delta t/\zeta$ with Δt being the time unit of the BD dynamics) and the temperature T in eqn. 1 give rise to a characteristic scale of the model. The two parameters specify the viscous force and the random force from a solvent and

are related to each other by eqn. 5. As is discussed in the accompanying paper II, the characteristic scale is the ‘persistence length’ of the chain. Actually, with $\zeta = 10$, $\Delta t = 1$ and $k_B T_{BD} = 2$ the mean distance between two connected beads (cl^* in eqn. 8) turns out to be about one six-th of the persistence length. Equating the latter to the persistence length of real DNA, the mean bead distance is roughly estimated as 10nm. Thus a bead in our BD model represents a portion of DNA consisting of a few tens of base-pairs. We may call our BD model with these values of parameters a bead-spring model of a semi-microscopic level. We will argue in paper III that the resultant chain dynamics, when it is coarse-grained in time and space, can be interpreted as dynamics of a charged, elastic string, and that the elasticity is due to the conformational entropy of the original, semi-microscopic chain.

A value of the electric field, $E = qE_b$ in eqn. 1, is reasonably converted to that of the experimental CFGE on DNA as discussed in Sec. III. By the procedure adopted there the conversion of the field scale is essentially governed by that of the length scale as expressed in eqn. 10. The parameter c in eqn. 10 has been adjusted for the chains with different sizes, but within the range that the results are compatible with the condition $\sqrt{2} \geq \langle l \rangle \geq 1$ where $\langle l \rangle$ is the mean distance of connected beads. Then the E -dependence of mobility μ_{BD} has been shown to agree semi-quantitatively with that of μ_{exp} once absolute magnitudes of the simulational and experimental μ ’s are adjusted at the vanishing field limit. This procedure fixes the conversion of the time scales of the simulation and the experiment. To our knowledge, such a semi-quantitative coincidence between the simulation and the experiment has not been so far achieved.

We have also demonstrated that our BD simulation on CFGE in fields stronger than a certain crossover value reproduces the quasi-periodic evolution of a chain whose characteristic aspects are quite similar to those experimentally observed. Actually, as proposed by Oana *et al.*¹⁵ based on their experimental observations, a whole sequence of time evolution of the chain can be divided into two branches according to behavior of the longer radius of the chain $R_l(t)$. One is the *deterministic* branch where $R_l(t)$ exhibits a distinct peak structure which corresponds to the elongation-contraction motion of the chain, and the other is the *stochastic* branch between two neighboring deterministic branches where the chain exhibits a rather compact form.

We have studied statistics of time durations of the two branches and have found that, as assumed by Oana *et al.*,¹⁵ those of the deterministic branch obey a Gaussian distribution

with mean D and variance Δ , while those of the stochastic branch do a Poisson distribution with the mean $1/\lambda$. Furthermore it is found that, in a fixed field, $D \propto M^1$, $\Delta \propto D$ and $1/\lambda \propto M^0$ for large M . This means that the deterministic branches become more dominant in a whole time sequence, and that the quasi-periodicity in the chain dynamics becomes more distinct (see Fig. 6) for larger M . These results, combined with the observation that the chain moves to the field direction by the distance proportional to M in each deterministic branch, yield $\mu \propto M^0$. This explains why CFGE cannot separate DNA according to their size L ($\propto M$) for large L . We have also checked that, for a fixed M which is sufficiently large, the deterministic branches become dominant as E is increased at least within the field range of our BD simulation (see Fig. 10).

By the present BD simulation a crossover between the regimes with and without elongation-contraction motions is expected to occur for a chain with $M \simeq 100$ in $E = 0.032$ and $a_{\text{gel}} = 17$. According to the length and field conversions described in Sec. III, this roughly corresponds to DNA of 2kbp in 1% agarose in $E_{\text{exp}} \simeq 4\text{V/cm}$. Experimentally, on the other hand, the direct measurement of the elongation-contraction motion by the fluorescent microscopy has been rather limited; the shortest is for T7 DNA (38kbp),²⁸ but mostly for T4 DNA (166kbp)¹⁵ and T2 DNA (164kbp).²⁸ But, there exist another experiment which strongly indicates the occurrence of the elongation-contraction motion, i.e., the antiresonance of mobility in the filed inversion gel electrophoresis. In such a experiment the antiresonance has been clearly observed for DNA with 9.42kbp.³¹ Therefore, our criterion mentioned above is considered to be a reasonable estimate for the elongation-contraction motion to occur.

Lastly two further comments are in order. In the whole parameter ranges we have examined, including the ones where the deterministic branches sufficiently dominate the stochastic ones, so-called 'hernias'-like configurations of a chain has been scarcely observed in the present BD simulation. This implies that they do not play a significant role in determining the saturation of μ at least for chains with moderate sizes we have studied. Another interesting observation has been already pointed out in Sec. IV B: the front end of a chain moves in the field direction with almost a constant rate (Fig. 4). In relation to this, we note that the average of $v_G(t)$ over the deterministic branches and that over the stochastic branches are found to almost coincide with each other in a whole range of field we have examined. We shall interpret these results in paper III where the details of the chain dynamics in the

deterministic branches is discussed.

Acknowledgments

The authors wish to thank M. Doi and Y. Masubuchi for useful discussions on their experimental and simulational results. The computation in this work has been done using the facilities of the Supercomputer Center, Institute for Solid State Physics, University of Tokyo, and those of the Computer Center of University of Tokyo.

References

- * Present address: Genomic Sciences Center, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa, 230-0045, Japan
- ¹ J.-L. Viovy, Rev. Mod. Phys. **72**, 813 (2000).
- ² B. Norden, C. Elvingston, M. Jonsson, and B. Åkerman, Quart. Rev. Biophys. **24**(2), 103 (1991).
- ³ D. C. Schwartz and C. R. Cantor, Cell **37**, 67 (1984).
- ⁴ D. Rodbard and A. Chrambach, Proc. Nat. Acad. Sci. **65**(4), 970 (1970).
- ⁵ A. Chrambach and D. Rodbard, Science **172**, 440 (1971).
- ⁶ M. Doi and S. F. Edwards, *The theory of polymer dynamics* (Clarendon Press, Oxford, 1986).
- ⁷ O. J. Lumpkin and B. H. Zimm, Biopolymers **21**, 2315 (1982).
- ⁸ O. J. Lumpkin, P. Déjardin, and B. H. Zimm, Biopolymers **24**, 1573 (1985).
- ⁹ G. W. Slater and J. Noolandi, Biopolymers **25**, 431 (1986).
- ¹⁰ N. C. Stellwagen, Biopolymers **24**, 2243 (1985).
- ¹¹ T. D. A.N. Semenov and J.-L. Viovy, Phys. Rev. E **51**, 1520 (1995).
- ¹² J. M. Deutsch, Science **240**, 922 (1988).
- ¹³ J. M. Deutsch and T. L. Madden, J. Chem. Phys. **90**(4), 2476 (1989).
- ¹⁴ Y. Masubuchi, H. Oana, K. Ono, M. Matsumoto, M. Doi, K. Minagawa, Y. Matsuzawa, and K. Yoshikawa, Macromolecules **26**, 5269 (1993).
- ¹⁵ H. Oana, Y. Masubuchi, M. Matsumoto, M. Doi, Y. Matsuzawa, and K. Yoshikawa, Macromolecules **27**(21), 6061 (1994).
- ¹⁶ S. B. Smith, L. Finzi, and C. Bustamante, Science **258**(13), 1122 (1992).
- ¹⁷ T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu, Science **268**(7), 83 (1995).
- ¹⁸ T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, Science **271**(29), 1835 (1996).
- ¹⁹ R. H. Austin, J. P. Brody, E. C. Cox, T. Duke, and W. Volkmuth, Physics Today **50**(2), 32 (1997).
- ²⁰ R. Azuma, preprint (2001).

- ²¹ P. G. de Gennes, J. Chem. Phys. **55**(2), 572 (1971).
- ²² R. Azuma and H. Takayama, in preparation (2001).
- ²³ T. D. C. Heller and J. Viovy, Biopolymers **34**, 249 (1994).
- ²⁴ W. D. Volkmuth, T. Duke, M. C. Wu, R. H. Austin, and A. Szabo, Phys. Rev. Lett. **72**(13), 2117 (1994).
- ²⁵ H. Hervet and C. P. Bean, Biopolymers **26**, 727 (1987).
- ²⁶ D. Holmes and N. Stellwagen, Electrophoresis **11**, 5 (1990).
- ²⁷ J. V. T. Duke and A. Semenov, Biopolymers **34**, 239 (1994).
- ²⁸ A. Larsson and B. Åkerman, Macromolecules **28**(13), 4441 (1995).
- ²⁹ M. Matsumoto and M. Doi, Molecular Simulation **12**, 219 (1994).
- ³⁰ Y. Masubuchi, H. Oana, and M. Matsumoto, in *Computational Physics as a New Frontier in Condensed Matter Research*, edited by H. Takayama, M. Tsukada, H. Shiba, F. Yonezawa, M. Imada, and Y. Okabe (The Physical Society of Japan, Tokyo, Japan, 1995), pp. 347–354.
- ³¹ Y. M. T. Kobayashi, M. Doi and M. Ogawa, Macromolecules **23**, 4480 (1990).

Figures

FIG. 1: A snapshot of $M = 10$ chain simulated in the BD method. The spherical objects express beads with a radius 0.5. There is no overlap between any pair of beads. The bar-like objects are inserted only to visualize bonding of neighboring beads.

FIG. 2: The scattering processes and the associated constraining forces to fulfill the constraint of eqn. 2 (a) and eqn. 3 (b).

FIG. 3: Field dependence of mobilities.

FIG. 4: A time evolution of projection of an $M = 240$ chain. $E = 0.032$ $a_{\text{gel}} = 17$. Solid line stands for locus of $i = 1$, and dashed line for $i = 240$.

FIG. 5: Time development of $R_l(t)$ and v_G under $E = 0.032$.

FIG. 6: The autocorrelation functions $C_{\text{RR}}(t)$. The dependence of period $4t_0$ on M is shown in the inset.

FIG. 7: (a) Distributions of peak widths $D_n = t_{\text{min:r}}^n - t_{\text{min:l}}^n$ and (b) duration between neighboring peak minima $1/\lambda_n = t_{\text{min:r}}^{n-1} - t_{\text{min:l}}^n$ obtained for $M = 240$, $E = 0.032$, $a_{\text{gel}} = 17$.

FIG. 8: Dependence of D and $1/\lambda$ on M .

FIG. 9: Ratios λD and Δ/D .

FIG. 10: Field dependence of DE , E/λ and $D\lambda$ of $M = 240$ chain.

FIG. 11: The averaged peak shape of $R_l(t)$. The abscissa is normalized by the period τ and the ordinate by the average of the peak height $\bar{R}_l(0)$.

FIG. 1

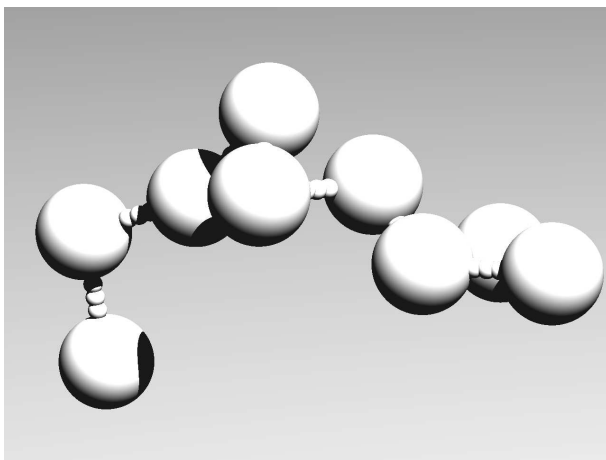


FIG. 2

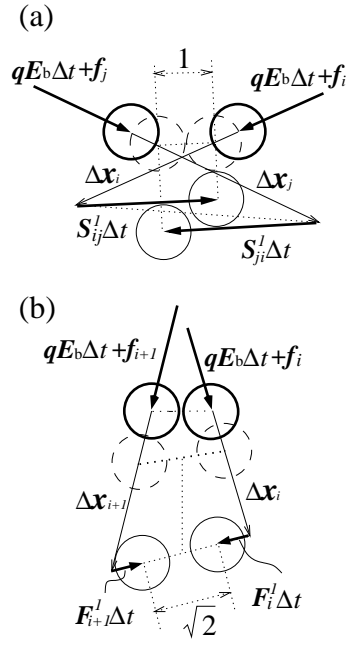


FIG. 3

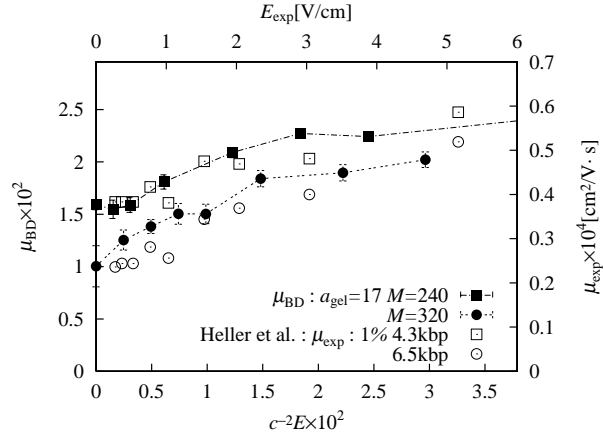


FIG. 4

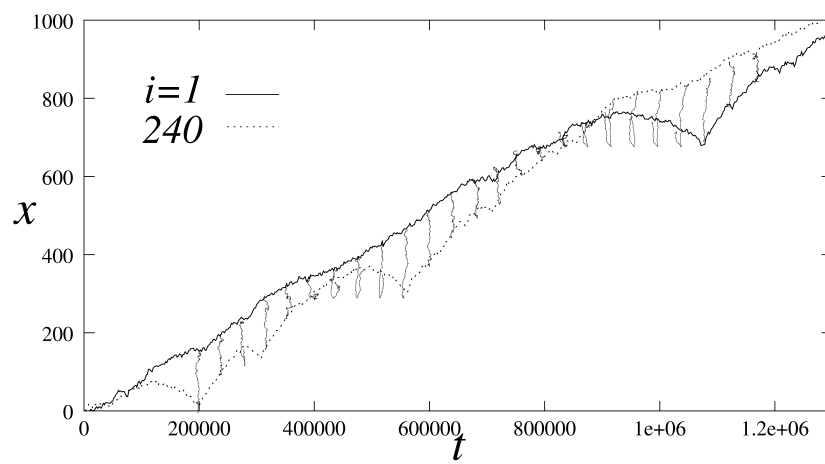


FIG. 5

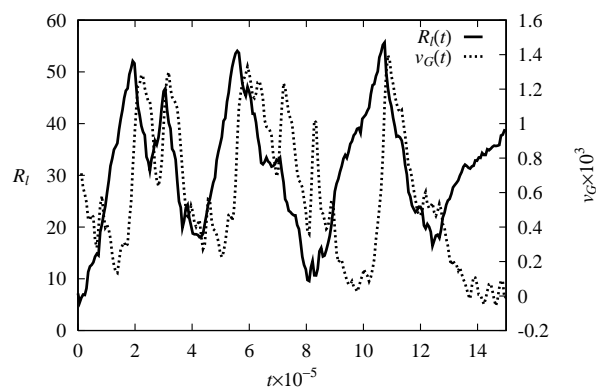


FIG. 6

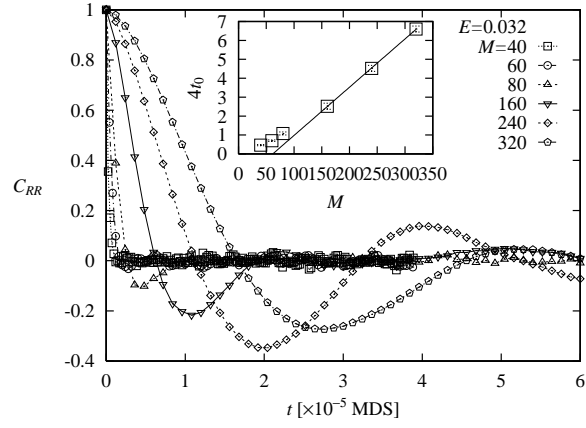


FIG. 7

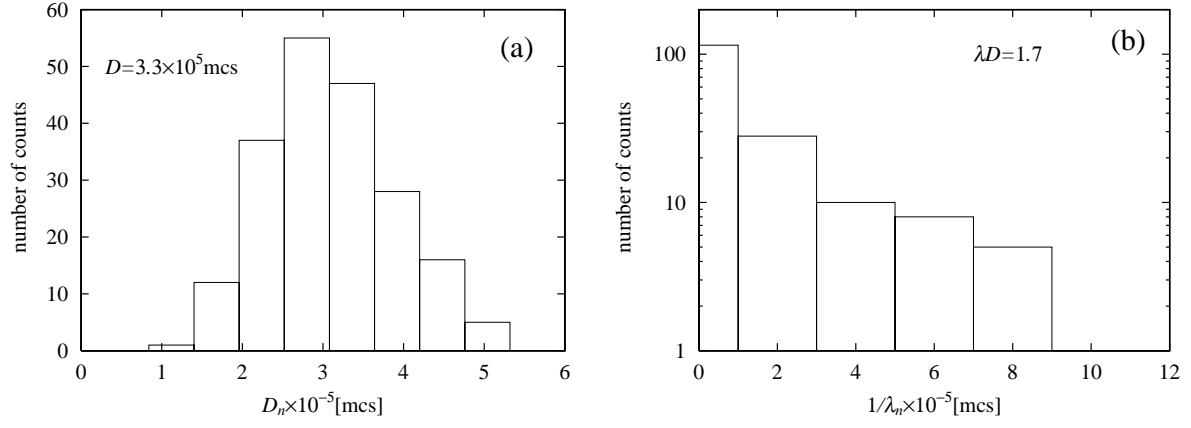


FIG. 8

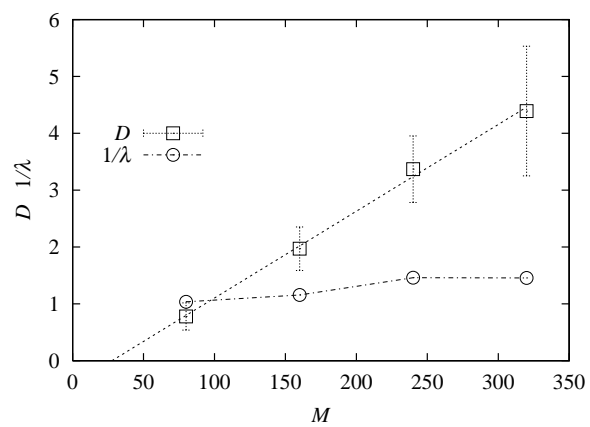


FIG. 9

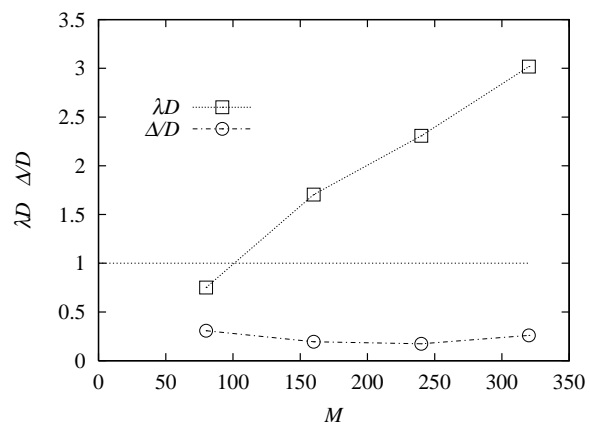


FIG. 10

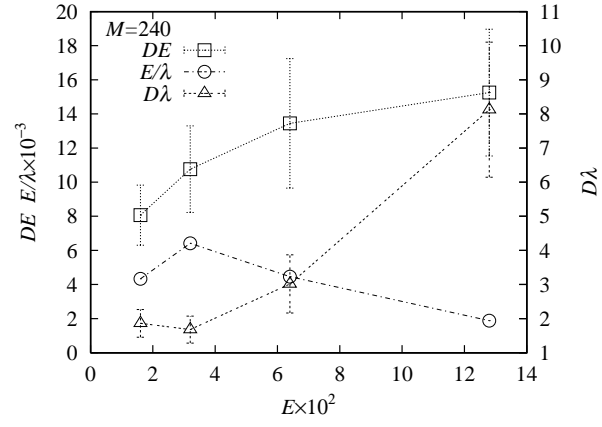


FIG. 11

